Flights Performance Analysis

A case study, leveraging scala and spark

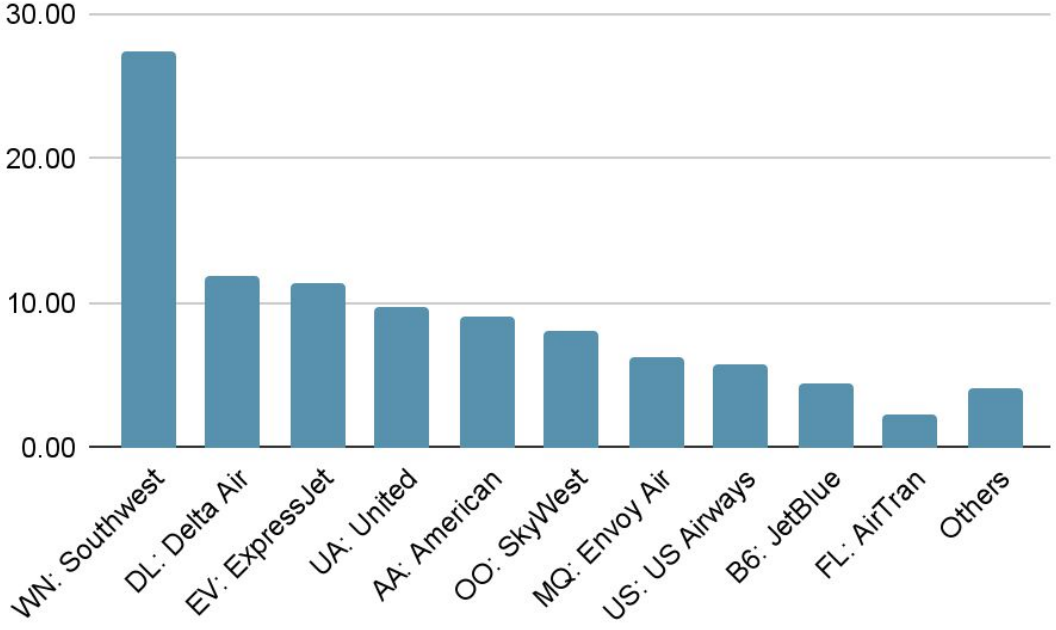
Batch A, Group 4

Abinaya N	CB.EN.U4AIE21001
Dilip P	CB.EN.U4AIE21009
Sanjana MCS	CB.EN.U4AIE21029
Siddharth	CB.EN.U4AIE21064

48%

Of flights are delayed, out of 450k flights we analysed

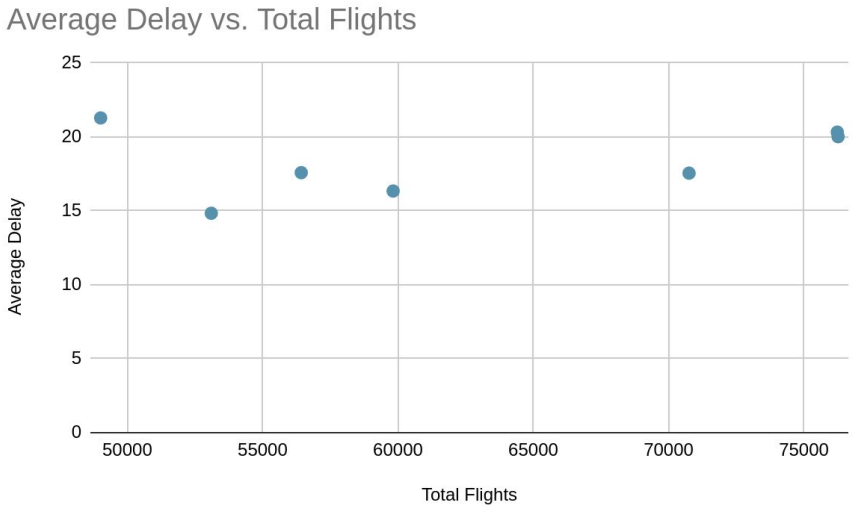
Southwest Airlines contributes the most to this delay



Carrier	Percentage of delay
WN: Southwest Airlines	27.42
DL: Delta Air Lines	11.86
EV: ExpressJet Airlines	11.46
UA: United Airlines	9.67
AA: American Airlines	9.06
OO: SkyWest Airlines	8.05
MQ: Envoy Air	6.22
US: US Airways	5.68
B6: JetBlue Airways	4.35
FL: AirTran Airways	2.21
Others	4.02

We found that, day of week, makes no impact in these delays. No weekend nothing..

Day of Week	Average Delay (in min)	Total Flights
Mon	16.3400331	59823
Tue	14.83586238	53103
Wed	17.54987916	70757
Thur	20.32430837	76233
Fri	20.01257228	76263
Sat	21.28247103	49016
Sun	17.58217874	56427
Total	18.38225541	441622



More the flights, more the delay…

Top delayed carriers and Top frequent carriers

Carrier	Average Delay
B6: JetBlue Airways	29.54
EV: ExpressJet Airlines	24.91
WN: Southwest Airlines	23.16
F9: Frontier Airlines	22.47
MQ: Envoy Air	20.13

Carrier	Total Flights
WN: Southwest Airlines	87161
DL: Delta Air Lines	56072
EV: ExpressJet Airlines	49626
OO: SkyWest Airlines	47924
AA: American Airlines	43852

The average delay is heavily impacted by, either more number of flights, or being some sort of international carrier.

This trend is also observed on per flight basis too

Correlation

The correlation between arrival delay and departure delay is

0.95

Flights that are delayed at departure are delayed even at arrival.
(roughly by the same factor)

Correlation between distance and speed of plane is

0.82

Longer the flight has to travel, faster its average speed.
(considering the cruise speed is the max, this verifies it)

Long traveling and high speed carriers

Carrier	Average Speed
AS: Alaska Airlines	364.44
F9: Frontier Airlines:	363.58
UA: United Airlines	362.82
VX: Virgin America	359.83
AA: American Airlines	347.33

Observe the correlation

Carrier	Average Distance
VX: Virgin America	1399.59
UA: United Airlines	1303.19
AS: Alaska Airlines	1196.58
AA: American Airlines	1066.89
B6: JetBlue Airways:	1066.86

We see that places traveling more distance in average tend to have higher average speed.

These planes are international carriers.

The average speed for all planes are 312.14 miles per hour (with std.dev of 83 mph)

Hardworking Planes

Tail Number	Carrier	Average Speed	Standard Deviation	Avg Delay	Number of Flights
N481HA	Hawaiian Airlines (HA)	205.72	38.23	2.5	356
N485HA	Hawaiian Airlines (HA)	199.29	37.38	1.03	350
N479HA	Hawaiian Airlines (HA)	201.46	36.38	1.75	347
N477HA	Hawaiian Airlines (HA)	198.63	35.47	1.15	329
N488HA	Hawaiian Airlines (HA)	201.38	37.55	1.34	329

(HA) Hawaiian airlines is a domestic flight, and this trend is same for all local flights.

Being a local flight, it travels less distance, getting less average speed.

On the other hand, local planes are mostly close to being on time (based on averages)

Distance of the flight vs delay

Distance	Avg Delay (min)
Very Short Haul (<100 miles)	16.20
Short Haul (100 to 500 miles)	17.22
Medium Haul (500 to 1000 miles)	19.18
Long Haul (>1000 miles)	19.04

This result directly corresponds to the frequency at which these planes operate. Domestic flights are usually more frequent (less distance) and they have lesser average delay

The Time of Day

Time of Day	Total Flights	Delayed Flights (%)	Average Delay (min)
Morning	37.47%	13.48%	25.65
Afternoon	38.37%	20.43%	36.09
Evening	20.63%	12.63%	51.38
Late Night	3.53%	1.03%	103.35

We see that there are very few flights in Late Night, but they contribute more to the delay in terms of time, although they only contribute 1% (to the 48%).

We also see that, mornings and afternoon are the busiest times for flight departure, followed by evening and late night flights

The Air Terminals

Terminal	Departure Count	Arrival Count	Total Delay (hrs)	Average Delay (mins)
ATL (Atlanta, Georgia, USA)	28420	28279	9692.8	20.46333568
DFW (Dallas/Fort Worth, Texas, USA)	22771	22724	4745.133333	12.50309604
LAX (Los Angeles, California, USA)	18051	18014	4455.95	14.8112016
ORD (Chicago, Illinois, USA)	17960	17874	8632.1	28.83775056
DEN (Denver, Colorado, USA)	17293	17175	5918.733333	20.53570809

These are the busiest air terminals. More planes contribute to more overall delays.

We also notice that DFW and LAX stations, have a comparatively better delays

Top flight routes

Top Routes	Num Flights	Avg Delay	Distance (miles)
San Francisco (SFO) to Los Angeles (LAX)	1080	33.17	337
Los Angeles (LAX) to San Francisco (SFO)	1074	31.97	337
Las Vegas (LAS) to Los Angeles (LAX)	1025	32.33	236
Los Angeles (LAX) to Las Vegas (LAS)	1021	29.58	236
John F. Kennedy International Airport (JFK) to Los Angeles (LAX)	868	41.34	2475
Los Angeles (LAX) to John F. Kennedy International Airport (JFK)	861	41.19	2475
Honolulu (HNL) to Kahului, Maui (OGG)	820	8.71	101
Kahului, Maui (OGG) to Honolulu (HNL)	820	7.29	101
Los Angeles (LAX) to Phoenix (PHX)	802	30.25	370
Phoenix (PHX) to Los Angeles (LAX)	798	25.02	370

Even here, we see that short flights have lower delay and longer flights have more delay!

The final Key takeaways

1. The major factor affecting the delay of the flights are time of departure and domestic nature.
2. We find that most frequent flights of the same planes are usually on time.
3. Strong correlation between arrival delay and departure, suggesting planes don’t go faster to cover up delays
4. Busy air terminals don’t affect the flight delays
5. Longer flights have higher average speed
6. The most impact to the average delay, is just the high number of flights. More the flights, more the chance of delay

Thank You